DOCUMENT RESUME

ED 149 318 CS 003 896

AUTHOR Cunningham, James W.; Tierney, Robert J.
TITLE Evaluating Cloze as a Measure of Cognitive Change Due to Reading.
PUB DATE Dec 77
NOTE 14p.; Paper presented at the Annual Meeting of the National Reading Conference (27th, New Orleans, Louisiana, December 1-3, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Cloze Procedure; *Cognitive Processes; *Comparative Analysis; Grade 7; Grade 8; Junior High Schools; Learning Processes; *Reading Comprehension; *Reading Research; Retention; Test Reliability; *Test Validity

ABSTRACT

This study compares the validity of three cloze procedures (traditional, limited, and least-major-constituent limited) when reading comprehension is defined as cognitive change. Over 400 seventh and eighth graders from public schools in North Carolina and Arizona were randomly assigned to one of the three cloze formats for either fiction or nonfiction passages, and to either an experimental or a control group. The experimental groups took the cloze test and one week later read the complete passage and took the cloze test again. Control group subjects followed the same procedures, but read complete passages that were unrelated to the cloze test. Results of the comparative analysis indicate that all of the cloze procedures were reliable measures, but that not all the cloze formats were consistently valid measures of reading comprehension defined as cognitive change. Two of the cloze formats (traditional and limited) were sensitive to the differential effects of reading, but only for the fiction passages. The study concludes that alternatives to cloze procedures should be sought when investigating reading comprehension as a cognitive process. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED149318

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

James W. Cunningham
Assistant Professor of C & I
University of North Carolina
307B Peabody Hall
Chapel Hill, N.C. 27514

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

James W. Cunningham
Robert J. Tierney

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Evaluating Cloze as a Measure of Cognitive Change Due to Reading

James W. Cunningham
University of North Carolina

Robert J. Tierney
University of Arizona

The investigators hold an essentially Piagetian view of reading comprehension, i.e. that reading comprehension is the differential modification of the reader's cognitive structures brought about by the reader's assimilation of and accommodation to concepts and relationships represented in print. Given such a view, it follows that measures of reading comprehension should reflect the modification of readers' cognitive structures. From this standpoint, the present study addressed an evaluation of the validity of cloze tests as measures of reading comprehension. Specifically, the question of concern was whether cloze scores reflect changes in cognitive structures due to reading.

CS203896

Background of Study

Since its inception (Taylor, 1953), the cloze testing procedure, in a variety of formats, has been claimed to be a reliable, objective measure of reading comprehension which correlates well with traditional reading comprehension measures (Taylor, 1957; Gallant, 1965; Bormuth, 1968; Rankin & Culhane, 1969). These psychometric strengths of cloze have lead to its use in a variety of research designs in the study of both readability (Klare, Simaiko, & Stolurow, 1972; Klare, 1974-75) and reading comprehension (Bormuth, 1969; Bickley, Ellington, & Bickley, 1970).

When reading comprehension, however, is defined as the change in the reader's schemata due to reading, as the investigators have, or as amount learned or information stored while reading, as Carver (1975) has, problems arise. Although Rankin and Dale (1969) advocated the use of cloze residual gain scores as a means of measuring learning through reading, Coleman and Miller (1968) concluded from their cloze data that group gain due to reading over and above group practice effect was negligible. While Rothkopf (1968) found that re-reading the unmutilated passage between repeated cloze testing improved scores relative to the repeated practice effect, Carver (1973; 1974) found cloze to be relatively insensitive to reading.

Because the literature is equivocal on the question of whether cloze scores reflect changes in schemata due to reading, the investigators felt that a more rigorous paradigm for evaluating the sensitivity to reading of cloze tests was called for. The study herein is the first

application of the paradigm to the specific problem of whether three cloze formats of two passages are sensitive to reading.

Method

Two passages, one fiction and one non-fiction, were selected at random from two textbooks published for use with developmental seventh-graders. Each of these passages was turned into three different cloze tests: a traditional nth cloze with n=5; a limited-cloze which is a traditional 5th cloze with the deleted words randomly ordered and put in columns on a separate sheet (Cunningham & Cunningham, 1978); and, a least-major-constituent limited-cloze in which every fifth least-major-constituent was deleted with all the deletions randomly ordered and put in columns on a separate sheet. Seventh- and eighth-graders from a public school in North Carolina were randomly assigned to one of the three cloze formats for the fiction passage. Seventh- and eighth-graders from a public school in Arizona were randomly assigned to one of the three cloze formats for the non-fiction passage. The subjects assigned to each cloze format of each passage were then randomly assigned to two groups, reading vs. buffer. The reading group took whichever cloze test was assigned to it, then a week later read the non-mutilated passage and took the same cloze test again. The buffer group followed the same steps except that the non-mutilated passage they read was not the passage from which the test had been constructed.

Data Analysis and Results

The first question to be answered was whether or not the data obtained from each of the six tests were reliable enough to warrant further analysis. All scores on the pre-test, for both the reading and buffer groups, were analyzed together to obtain a Kuder-Richardson 20 reliability coefficient for each of the six tests. In addition, the correlation between pre- and post-test scores for the buffer group in each test condition served as a test-retest reliability coefficient for that test. Results are presented in Table 1.

Insert Table 1 about here

With the possible exception of the non-fiction regular-cloze test, all test data appeared to be reliable, some extraordinarily so. The decision was made to give the non-fiction regular-cloze data the benefit of the doubt and the analysis of data from all six test conditions was continued.

The second question to be answered was whether or not any or all of the six cloze tests were sensitive to reading, i.e. would the difference between pre- and post-test scores be significantly greater for the group reading the passage in between than for the group reading the buffer in between. It would, of course, be impossible to infer cognitive change due to reading when there is no change due to reading over practice effect. To answer this question, a two-way analysis of variance with a repeated measure as the second factor was performed on the data for each of the six tests. A significant interaction for a test condition (interpreted

by observation of cell means) would support that there were greater gains pre- to post- for the reading group than for the buffer group. An alpha level of .05 was selected for all testing. Results are presented in Table 2.

Insert Table 2 about here

Five of the six interactions were significant (with cell-mean differences in the proper directions), i.e. both regular-cloze tests, both limited-cloze tests, and the non-fiction LMC limited-cloze had significant sensitivity to reading over and above group practice effect.

The third question to be answered was whether or not any or all of the five cloze tests shown (above) to be sensitive to reading were sensitive to individual differences in subjects' ability to gain information by reading the intact passage prior to the post-test. The logic used to answer this question was as follows:

(a) Some or all of the five cloze tests could have been sensitive to subjects reading the passage between pre- and post-testing because the test deleted a few items which were not highly redundant yet which were memorable after reading the non-mutilated passage. These few items would tend to inflate everyone's post-test score in the reading group causing a greater difference between pre- and post-test scores. While this difference is a sensitivity-to-reading effect, it would tend to raise each subject's score equally. Certainly, a true measure of information-gain during reading would be expected to be sensitive to individual differences.

(b) Although the theoretical ceiling for a correlation between pre- and post-test scores is the reliability of the test (Thorndike & Hagen, 1969), the test-retest correlation between pre- and post-scores for any group reading the passage in between would be expected to approach the test's reliability if the effect due to reading approached a constant.

To determine whether or not any or all of the five cloze tests had been sensitive to reading by merely raising each subject's pre-test score by what approached a constant amount, the test-retest coefficient for the reading group in each of the five test conditions was tested for the significance of the difference between it and the test-retest reliability of the test. Results are in Table 3.

Insert Table 3 about here

Surprisingly, three of the five reading-group test-retest coefficients were not significantly lower than the test-retest reliability of the test. These results suggest that the non-fiction regular-cloze, the non-fiction limited-cloze, and the non-fiction LMC limited-cloze had a significant effect due to reading only because reading added a near-constant amount to each subject's pre-test score. An examination of the item means for the three tests revealed that, for the non-fiction regular cloze, just four items out of the 40 accounted for 51% of the estimated total reading effect and 85% of the estimated sensitivity-to-reading effect; for the non-fiction limited-cloze, seven items out of the 40 accounted for 35% of the estimated total reading effect and 52% of the estimated sensitivity-to-reading effect; and, for the non-fiction

IMC limited-cloze, five items out of the 25 accounted for 48% of the estimated total reading effect and 52% of the estimated sensitivity-to-reading effect.

At last, only two of the six cloze tests passed the three criteria of reliability, sensitivity to reading, and sensitivity to individual differences in reading. These two tests were the fiction regular-cloze and the fiction limited-cloze.

Discussion

Obviously, the study is limited by the specific passages used, the specific cloze formats used, and the specific subjects tested. Also the selection of alpha level was somewhat arbitrary. If an alpha level of .01 had been selected, only one test--the fiction regular-cloze--would have passed the criteria of whether cloze scores reflect changes in cognitive structures due to reading.

Assuming that these limitations have not caused unusual results, the data have some implications for using cloze scores to investigate reading comprehension as the modification of readers' schemata due to reading. When using cloze scores for such a purpose, the following steps should be taken:

1. Some rigorous paradigm, such as the one above, should be used at a pilot stage to select specific cloze tests which are reliable, sensitive to reading, and sensitive to individual differences in reading for the subjects to be used in the study.

2. Once a specific test has passed these three criteria, it should be used in a Soloman Four-Group Design or the equivalent where the unit of experimentation is the class or group rather than the individual. This process would allow the mean for each group to be partitioned into prior-knowledge effect, practice effect, sensitivity-to-reading effect, and the various interaction effects.
3. The sensitivity-to-reading effect partitioned from each group mean could be used as a dependent or criterion variable in research on reading comprehension cognitive processes.

It must seem that such a process would be incredibly costly in both time and money and access to subjects. Surely there is some less costly alternative to cloze for investigating comprehension as a cognitive process.

References

Bickley, A. C., Ellington, B. J., & Bickley, R. T. The cloze procedure: A conspectus. Journal of Reading Behavior, 1970, 2, 232-249.

Bormuth, J. R. Cloze test readability: Criterion-references scores. Journal of Educational Measurement, 1968, 5, 189-196.

Bormuth, J. R. Factor validity of cloze tests as measures of reading comprehension ability. Reading Research Quarterly, 1969, 4, 358-365.

Carver, R. P. Understanding, information-processing, and learning from prose materials. Journal of Educational Psychology, 1973, 64, 76-84.

Carver, R. P. Measuring the primary effect of reading: Reading-storage technique, understanding judgments, and cloze. Journal of Reading Behavior, 1974, 6, 249-274.

Carver, R. P. Revised procedures for developing reading-input materials and reading-storage tests. Journal of Reading Behavior, 1975, 7, 155-172.

Coleman, E. B., & Miller, G. R. A measure of information gained during prose learning. Reading Research Quarterly, 1968, 3, 369-386.

Cunningham, J. W., & Cunningham, P. M. Validating a limited-cloze procedure. Journal of Reading Behavior, 1978, in press.

Gallant, R. Use of cloze tests as a measure of readability in the primary grades. In J. A. Figurel (Ed.), Reading and inquiry. Newark, Delaware: International Reading Association, 1965.

Klare, G. R. Assessing readability. Reading Research Quarterly, 1974-75, 10, 62-102.

Klare, G. R., Simaiko, H. W., & Stolurow, L. M. The cloze procedure: A convenient readability test for training materials and translations. International Review of Applied Psychology, 1972, 21(2), 77-106.

Rankin, E. F., & Culhane, J. Comparable cloze and multiple-choice comprehension test scores. Journal of Reading, 1969, 13, 193-198.

Rankin, E. F., & Dale, L. H. Cloze residual gain - A technique for measuring learning through reading. In G. B. Schick & M. M. May (Eds.), The psychology of reading behavior, Eighteenth Yearbook, National Reading Conference. Milwaukee: National Reading Conference, 1969.

Rothkopf, E. Z. Textual constraint as function of repeated inspection. Journal of Educational Psychology, 1968, 59, 20-25.

Taylor, W. Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-433.

Taylor, W. L. "Cloze" readability scores as indices of individual differences in comprehension and aptitude. Journal of Applied Psychology, 1957, 41, 19-26.

Thorndike, R. L., & Hagen, E. Measurement and evaluation in psychology and education (3rd ed.). New York: John Wiley & Sons, 1969.

Table 1

KR 20 and Test-Retest Reliabilities

	Cloze Format					
	Cloze		Limited-Cloze		LMC Limited-Cloze	
Passage	KR 20	Test/Retest	KR 20	Test/Retest	KR 20	Test/Retest
Fiction	.92	.96	.95	.94	.93	.82
Non-Fiction	.67	.67	.90	.85	.86	.80

Table 2

Two-Way Anovas with Repeated Measure on One Factor

Significance of Interaction

	Cloze Format		
Passage	Cloze	Limited-Cloze	LMC Limited-Cloze
Fiction	$p<.001$ (n=81)	$p<.001$ (n=81)	$p<.56$ (N.S.) (n=82)
Non-Fiction	$p<.01$ (n=63)	$p<.02$ (n=82)	$p<.01$ (n=68)

N.S. = $p>.05$

Table 3

One-Tailed Tests of Significance of Difference between Reading-Group Test-Retest Coefficients and Test Reliabilities Using Fisher's Z_r Transformation

	Cloze Format		
Passage	Cloze	Limited-Cloze	LMC Limited-Cloze
Fiction	$p < .01$	$p < .04$	- - - - - -
Non-Fiction	$p < .18$ (N.S.)	$p < .99$ (N.S.)	$p < .35$ (N.S.)

N.S. = $p > .05$